

OIPE

#2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/848,915

DATE: 06/14/2001

TIME: 10:24:22

Input Set : N:\Crf3\RULE60\09848915.txt

Output Set: N:\CRF3\06142001\I848915.raw

SEQUENCE LISTING

C--> 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Hillman, Jennifer L.

8 Goli, Surya K.

C--> 10 (ii) TITLE OF INVENTION: NOVEL TUMORIGENESIS PROTEIN

12 (iii) NUMBER OF SEQUENCES: 3

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

16 (B) STREET: 3174 Porter Drive

17 (C) CITY: Palo Alto

18 (D) STATE: CA

19 (E) COUNTRY: USA

20 (F) ZIP: 94304

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Diskette

24 (B) COMPUTER: IBM Compatible

25 (C) OPERATING SYSTEM: DOS

26 (D) SOFTWARE: FastSEQ for Windows Version 2.0

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/848,915

C--> 30 (B) FILING DATE: 04-May-2001

31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/822,260

35 (B) FILING DATE:

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Billings, Lucy J.

39 (B) REGISTRATION NUMBER: 36,749

40 (C) REFERENCE/DOCKET NUMBER: PF-0247 US

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 415-855-0555

44 (B) TELEFAX: 415-845-4166

45 (C) TELEX:

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 195 amino acids

52 (B) TYPE: amino acid

53 (C) STRANDEDNESS: single

54 (D) TOPOLOGY: linear

56 (vii) IMMEDIATE SOURCE:

57 (A) LIBRARY: UTRSN02

58 (B) CLONE: 2267574

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

62 Met Glu Leu Ser Glu Ser Val Gln Lys Gly Phe Gln Met Leu Ala Asp

63 1 5 10 15

64 Pro Arg Ser Phe Asp Ser Asn Ala Phe Thr Leu Leu Leu Arg Ala Ala

65 20 25 30

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/848,915

DATE: 06/14/2001

TIME: 10:24:22

Input Set : N:\Crf3\RULE60\09848915.txt

Output Set: N:\CRF3\06142001\I848915.raw

```

66 Phe Gln Ser Leu Leu Asp Ala Gln Ala Asp Glu Ala Val Leu Asp His
67      35      40      45
68 Pro Asp Leu Lys His Ile Asp Pro Val Val Leu Lys His Cys His Ala
69      50      55      60
70 Ala Ala Ala Thr Tyr Ile Leu Glu Ala Gly Lys His Arg Ala Asp Lys
71      65      70      75      80
72 Ser Thr Leu Ser Thr Tyr Leu Glu Asp Cys Lys Phe Asp Arg Glu Arg
73      85      90      95
74 Ile Glu Leu Phe Cys Thr Glu Tyr Gln Asn Asn Lys Asn Ser Leu Glu
75      100      105      110
76 Ile Leu Leu Gly Ser Ile Gly Arg Ser Leu Pro His Ile Thr Asp Val
77      115      120      125
78 Ser Trp Arg Leu Glu Tyr Gln Ile Lys Thr Asn Gln Leu His Arg Met
79      130      135      140
80 Tyr Arg Pro Ala Tyr Leu Val Thr Leu Ser Val Gln Asn Thr Asp Ser
81      145      150      155      160
82 Pro Ser Tyr Pro Glu Ile Ser Phe Ser Cys Ser Met Glu Gln Leu Gln
83      165      170      175
84 Asp Leu Val Gly Lys Leu Lys Asp Ala Ser Lys Ser Leu Glu Arg Ala
85      180      185      190
86 Thr Gln Leu
87      195

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 751 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: UTRSNOT02
- (B) CLONE: 2267574

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

101 CGAAGTCACG GCGCGCTCAC AATGGAGCTC TCGGAGTCTG TGCAGAAAGG CTTCCAGATG      60
102 CTGGCGGATC CCCGCTCCTT CGACTCCAAC GCCTTCACGC TTCTCCTCCG GGCGGCATTC      120
103 CAGAGTCTGC TGGACGCCCA GGCGGACGAG GCCGTGTTAG ATCATCCAGA CTTGAAACAT      180
104 ATCGACCCAG TGGTTTTTAA ACATTGTCAT GCAGCAGCTG CAACTTACAT ACTAGAGGCA      240
105 GGAAAGCACC GAGCTGACAA GTCAACTCTA AGCACTTATC TAGAAGACTG TAAATTTGAC      300
106 AGAGAGCGAA TAGAACTGTT TTGCACGGAA TATCAGAATA ATAAGAATTC CCTAGAAATC      360
107 CTAAGGGGAA GTATAGGCAG ATCTCTCCCT CATATAACGG ATGTTTCTTG GCGCTTGGAA      420
108 TATCAGATAA AGACCAATCA ACTTCATAGG ATGTACAGAC CTGCATATTT GGTGACCTTA      480
109 AGTGTACAGA AACTGATTC CCCATCCTAT CCAGAGATTA GTTTTAGTTG CAGCATGGAA      540
110 CAATTACAGG ACTTGGTGGG GAAACTTAAA GATGCTTCGA AAAGCCTGGA AAGAGCAACT      600
111 CAGTTGTAAC TTGGGGAAGT TAACGATCCG CCCGAGTGCA GAGGAAAACC AGAAACGCCT      660
112 TGCCTTCAGC TGAACCACCG TTTGTGCGAG CTGGATGTCC TTTTCAGTAG AAAAGAATTT      720
113 TCCTTTTGAA TTTATACCAT TCANCAATTT T
114
115
116
117

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 amino acids
- (B) TYPE: amino acid

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/848,915

DATE: 06/14/2001

TIME: 10:24:22

Input Set : N:\Crf3\RULE60\09848915.txt

Output Set: N:\CRF3\06142001\I848915.raw

```

122      (C) STRANDEDNESS: single
123      (D) TOPOLOGY: linear
125      (vii) IMMEDIATE SOURCE:
126          (A) LIBRARY: GenBank
127          (B) CLONE: 265569
129      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
131 Met Glu Leu Ser Glu Ser Val Gln Arg Gly Ile Gln Thr Leu Ala Asp
132   1          5          10          15
133 Pro Gly Ser Phe Asp Ser Asn Ala Phe Ala Leu Leu Leu Arg Ala Ala
134          20          25          30
135 Phe Gln Ser Leu Leu Asp Ala Arg Ala Asp Glu Ala Ala Leu Asp His
136          35          40          45
137 Pro Tyr Leu Lys Gln Ile Asp Pro Val Val Leu Lys His Cys His Ala
138          50          55          60
139 Ala Ala Ala Thr Cys Ile Leu Glu Ala Gly Lys His Gln Val Asp Lys
140          65          70          75          80
141 Ser Thr Leu Ser Thr Tyr Leu Glu Asp Cys Lys Phe Asp Arg Glu Arg
142          85          90          95
143 Ile Glu Leu Phe Cys Thr Glu Tyr Gln Asn Asn Lys Asn Ser Leu Glu
144          100          105          110
145 Thr Leu Leu Gly Ser Ile Gly Arg Ser Leu Pro His Ile Thr Asp Val
146          115          120          125
147 Ser Trp Arg Leu Glu Tyr Gln Ile Lys Thr Asn Gln Leu His Lys Met
148          130          135          140
149 Tyr Arg Pro Gly Tyr Leu Val Thr Leu Asn Val Glu Asn Asn Asp Ser
150          145          150          155          160
151 Gln Ser Tyr Pro Glu Ile Asn Phe Ser Cys Asn Met Glu Gln Leu Gln
152          165          170          175
153 Asp Leu Val Gly Lys Leu Lys Asp Ala Ser Lys Ser Leu Glu Arg Ala
154          180          185          190
155 Thr Gln Leu
156          195

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/848,915

DATE: 06/14/2001

TIME: 10:24:23

Input Set : N:\Crf3\RULE60\09848915.txt

Output Set: N:\CRF3\06142001\I848915.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:10 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]